Last time we considered the IVP for the heat equation on the whole line

$$\begin{cases} u_t - ku_{xx} = 0 \quad (-\infty < x < \infty, 0 < t < \infty), \\ u(x,0) = \phi(x), \end{cases} \tag{10.1}$$

and derived the solution formula

$$u(x,t) = \int_{-\infty}^{\infty} S(x-y,t)\phi(y)\,dy, \qquad \text{for} \quad t > 0, \tag{10.2}$$

where $S(x,t)$ is the heat kernel,

$$S(x,t) = \frac{1}{\sqrt{4\pi kt}} e^{-x^2/4kt}. \tag{10.3}$$

Substituting this expression into (10.2), we can rewrite the solution as

$$u(x,t) = \frac{1}{\sqrt{4\pi kt}} \int_{-\infty}^{\infty} e^{-(x-y)^2/4kt}\phi(y)\,dy, \qquad \text{for} \quad t > 0. \tag{10.4}$$

Recall that to derive the solution formula we first considered the heat IVP with the following particular initial data

$$Q(x,0) = H(x) = \begin{cases} 1, & x > 0, \\ 0, & x < 0. \end{cases} \tag{10.5}$$

Then using dilation invariance of the Heaviside step function $H(x)$, and the uniqueness of solutions to the heat IVP on the whole line, we deduced that Q depends only on the ratio $x/\sqrt{t}$, which lead to a reduction of the heat equation to an ODE. Solving the ODE and checking the initial condition (10.5), we arrived at the following explicit solution

$$Q(x,t) = \frac{1}{2} + \frac{1}{\sqrt{\pi}} \int_0^{x/\sqrt{4kt}} e^{-p^2}\,dp, \qquad \text{for} \quad t > 0. \tag{10.6}$$

The heat kernel $S(x,t)$ was then defined as the spatial derivative of this particular solution $Q(x,t)$, i.e.

$$S(x,t) = \frac{\partial Q}{\partial x}(x,t), \tag{10.7}$$

and hence it also solves the heat equation by the differentiation property.

The key to understanding the solution formula (10.2) is to understand the behavior of the heat kernel $S(x,t)$. To this end some technical machinery is needed, which we develop next.

10.1 Dirac delta function

Notice that, due to the discontinuity in the initial data of Q, the derivative $Q_x(x,t)$, which we used in the definition of the function S in (10.7), is not defined in the traditional sense when $t = 0$. So how can one make sense of this derivative, and what is the initial data for $S(x,t)$?

It is not difficult to see that the problem is at the point $x = 0$. Indeed, using that $Q(x,0) = H(x)$ is constant for any $x \neq 0$, we will have $S(x,0) = 0$ for all x different from zero. However, $H(x)$ has a jump discontinuity at $x = 0$, as is seen in Figure 10.1, and one can imagine that at this point the rate of growth of H is infinite. Then the "derivative"

$$\delta(x) = H'(x) \tag{10.8}$$

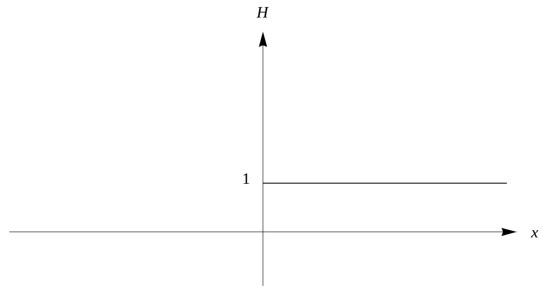

Figure 10.1: The graph of the Heaviside step function.

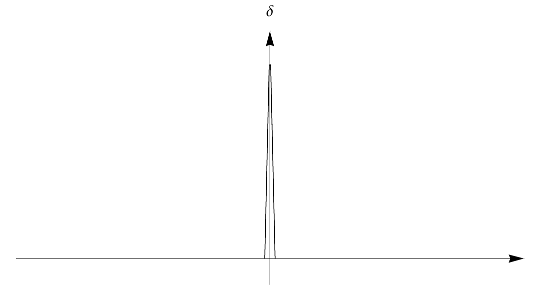

Figure 10.2: The sketch of the Dirac δ function.

is zero everywhere, except at $x = 0$, where it has a spike of zero width and infinite height. Refer to Figure 10.2 below for an intuitive sketch of the graph of δ. Of course, δ is not a function in the traditional sense, but is rather a *generalized function*, or *distribution*. Unlike regular functions, which are characterized by their finite values at every point in their domains, distributions are characterized by how they act on regular functions.

To make this rigorous, we define the set of *test functions* $\mathcal{D} = C_c^\infty$, the elements of which are smooth functions with compact support. So $\phi \in \mathcal{D}$, if and only if ϕ has continuous derivatives of any order $k \in \mathbb{N}$, and the closure of the support of ϕ,

$$\operatorname{supp}(\phi) = \{x \in \mathbb{R} \mid \phi(x) \neq 0\},$$

is compact. Recall that compact sets in $\mathbb{R}$ are those that are closed and bounded. In particular for any test function ϕ there is a rectangle $[-R, R]$, outside of which ϕ vanishes. Notice that derivatives of test functions are also test functions, as are sums, scalar multiples and products of test functions.

Distributions are continuous linear *functionals* on $\mathcal{D}$, that is, they are continuous linear maps from $\mathcal{D}$ to the real numbers $\mathbb{R}$. Notice that for any regular function f, we can define the functional

$$f[\phi] = \int_{-\infty}^{\infty} f(x)\phi(x)\,dx, \tag{10.9}$$

which makes f into a distribution, since to every $\phi \in \mathcal{D}$ it assigns the number $\int_{-\infty}^{\infty} f(x)\phi(x)\,dx$. This integral will converge under very weak conditions on f ($f \in L^1_{loc}$), due to the compact support of ϕ. In particular, f can certainly have jump discontinuities. Notice that we committed an abuse of notation to identify the distribution associated with f by the same letter f. The particular notion in which we use the function will be clear from the context.

One can also define the *distributional derivative* of f to be the distribution, which acts on the test functions as follows

$$f'[\phi] = -\int_{-\infty}^{\infty} f(x)\phi'(x)\,dx.$$

Notice that integration by parts and the compact support of test functions makes this definition consistent with the regular derivative for differentiable functions (check that the distribution formed as in (10.9) by the derivative of f coincides with the distributional derivative of f).

We can also apply the notion of the distributional derivative to the Heaviside step function $H(x)$, and think of the definition (10.8) in the sense of distributional derivatives. Let us now compute how δ, called the *Dirac delta function*, acts on test functions. By the definition of the distributional derivative,

$$\delta[\phi] = -\int_{-\infty}^{\infty} H(x)\phi'(x)\,dx.$$

Recalling the definition of $H(x)$ in (10.5), we have that

$$\delta[\phi] = -\int_{0}^{\infty} \phi'(x)\,dx = -\phi(x)\Big|_0^\infty = \phi(0). \tag{10.10}$$

Thus, the Dirac delta function maps test functions to their values at $x=0$. We can make a translation in the x variable, and define $\delta(x-y) = H'(x-y)$, i.e. $\delta(x-y)$ is the distributional derivative of the distribution formed by the function $H(x-y)$. Then it is not difficult to see that $\delta(x-y)[\phi] = \phi(y)$. That is, $\delta(x-y)$ maps test functions to their values at y. We will make the abuse of notation mentioned above, and write this as

$$\int_{-\infty}^{\infty} \delta(x-y)\phi(x)\,dx = \phi(y).$$

We also note that $\delta(x-y) = \delta(y-x)$, since δ is even, if we think of it as a regular function with a centered spike (one can prove this from the definition of δ as a distribution).

Using these new notions, we can make sense of the initial data for $S(x,y)$. Indeed,

$$S(x,0) = \delta(x). \tag{10.11}$$

Since the initial data is a distribution, one then thinks of the equation to be in the sense of distributions as well, that is, treat the derivatives appearing in the equation as distributional derivatives. This requires the generalization of the idea of a distribution to two dimensions. We call this type of solutions *weak solutions* (recall the solutions of the wave equation with discontinuous data). Thus $S(x,t)$ is a week solution of the heat equation, if

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(x,t)[\phi_t(x,t) - k\phi_{xx}(x,t)]\,dxdt = 0,$$

for any test function ϕ of two variables. This means that the distribution $(\partial_t - k\partial_x^2)S$, with the derivatives taken in the distributional sense, is the zero distribution. Notice that the weak solution $S(x,t)$ arising from the initial data (10.11) has the form (10.3), which is an infinitely differentiable function of x and t. This is in stark contrast to the case of the wave equation, where, as we have seen in the examples, the discontinuity of the initial data is preserved in time.

Having the δ function in our arsenal of tools, we can now give an alternate proof that (10.2) satisfies the initial conditions of (10.1). Directly plugging in $t=0$ into (10.2), which we are now allowed to do by treating it as a distribution, and using (10.11), we get

$$u(x,0) = \int_{-\infty}^{\infty} \delta(x-y)\phi(y)\,dy = \phi(x).$$

10.2 Interpretation of the solution

Let us look at the solution (10.4) in detail, and try to understand how the heat kernel $S(x,t)$ propagates the initial data $\phi(x)$. Notice that $S(x,t)$, given by (10.3), is a well-defined function of (x,t) for any $t>0$. Moreover, $S(x,t)$ is positive, is even in the x variable, and for a fixed t has a bell-shaped graph. In general, the function

$$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

is called *Gaussian function* or *Gaussian*. In the probability theory, it gives the density of the normal distribution with mean μ and standard deviation σ. The graph of the Gaussian is a bell-curve with its peak of height $1/\sqrt{2\pi\sigma^2}$ at $x=\mu$ and the width of the bell at mid-height roughly equal to 2σ. Thus, for some fixed time t the height of $S(x,t)$ at its peak $x=0$ is $\frac{1}{\sqrt{4\pi kt}}$, which decays as t grows.

Notice that as $t \to 0+$, the height of the peak becomes arbitrarily large, and the width of the bell-curve, $\sqrt{2kt}$ goes to zero. This, of course, is expected, since $S(x,t)$ has the initial data (10.11). One can think of $S(x,t)$ as the temperature distribution at time t that arises from the initial distribution given by the Dirac delta function. With passing time the highest temperature at $x=0$ gets gradually transferred to the other points of the rod. It also makes sense, that points closer to $x=0$ will have higher temperature than those farther away. Graphs of $S(x,t)$ for three different times are sketched in Figure 10.3 below.

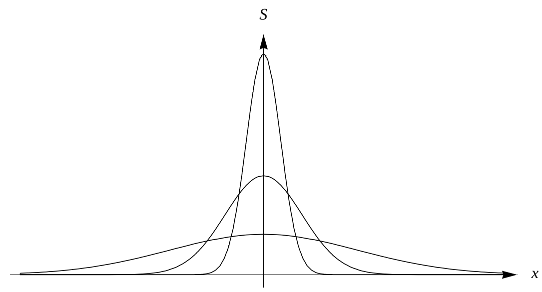

Figure 10.3: The graphs of the heat kernel at different times.

From the initial condition (10.11), we see that initially the temperature at every point $x \neq 0$ is zero, but $S(x,t) > 0$ for any x and $t > 0$. This means that heat is instantaneously transferred to all points of the rod (closer points get more heat), so the speed of heat conduction is infinite. Compare this to the finite speed of propagation for the wave equation. One can also compute the area below the graph of $S(x,t)$ at any time $t > 0$ to get

$$\int_{-\infty}^{\infty} S(x,t)\, dx = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-p^2}\, dp = 1,$$

where we used the change of variables $p = x/\sqrt{4kt}$. At $t = 0$, we have

$$\int_{-\infty}^{\infty} S(x,0)\, dx = \int_{-\infty}^{\infty} \delta(x)\, dx = 1,$$

where we think of the last integral as the δ distribution applied to the constant function 1 (more precisely, a test function that is equal to 1 in some open interval around $x = 0$). This shows that the area below the graph of $S(x,t)$ is preserved in time and is equal to 1, so for any fixed time $t \geq 0$, $S(x,t)$ can be thought of as a probability density function. At time $t = 0$ its the probability density that assigns probability 1 to the point $x = 0$, as was seen in (10.10), and for times $t > 0$ it is a normal distribution with mean $x = 0$ and standard deviation $\sigma = \sqrt{2kt}$ that grows with time. As we mentioned earlier, $S(x,t)$ is smooth, in spite of having a discontinuous initial data. We will see in the next lecture that this is true for any solution of the heat IVP (10.1) with general initial data.

We now look at the solution (10.4) with general data $\phi(x)$. First, notice that the integrand in (10.2),

$$S(x-y,t)\phi(y),$$

measures the effect of $\phi(y)$ (the initial temperature at the point y) felt at the point x at some later time t. The source function $S(x-y,t)$, which has its peak precisely at y, *weights* the contribution of $\phi(y)$ according to the distance of y from x and the elapsed time t.

Since the value of $u(x,t)$ (temperature at the point x at time t) is the total sum of contributions from the initial temperature at all points y, we have the formal sum

$$u(x,t) \approx \sum_{y} S(x-y,t)\phi(y),$$

which in the limit gives formula (10.2). So, the heat kernel $S(x,t)$ gives a way of propagating the initial data ϕ to later times. Of course the contribution from a point y_1 closer to x has a bigger weight $S(x-y_1,t)$, than the contribution from a point y_2 farther away, which gets weighted by $S(x-y_2,t)$.

The function $S(x,t)$ appears in various other physical situations. For example in the random (Brownian) motion of a particle in one dimension. If the probability of finding the particle at position x initially is given by the density function $\phi(x)$, then the density defining the probability of finding the particle at position x at time t is given by the same formula (10.2).

Example 10.1. Solve the heat equation with the initial condition $u(x,0) = e^x$.
Using the solution formula (10.4), we have

$$u(x,t) = \frac{1}{\sqrt{4\pi kt}} \int_{-\infty}^{\infty} e^{-(x-y)^2/4kt} e^y \, dy = \frac{1}{\sqrt{4\pi kt}} \int_{-\infty}^{\infty} e^{[-x^2+2xy-y^2+4kty]/4kt} \, dy$$

We can complete the squares in the numerator of the exponent, writing it as

$$\begin{aligned}\frac{-x^2+2xy-y^2+4kty}{4kt} &= \frac{-x^2+2(x+2kt)y-y^2}{4kt} \\ &= \frac{-(y-2kt-x)^2+4ktx+4k^2t^2}{4kt} = -\left(\frac{y-2kt-x}{\sqrt{4kt}}\right)^2 + x + kt.\end{aligned}$$

We then have

$$u(x,t) = \frac{1}{\sqrt{4\pi kt}} \int_{-\infty}^{\infty} e^{x+kt} e^{[(y-2kt-x)/\sqrt{4kt}]^2} \, dy = \frac{1}{\sqrt{\pi}} e^{x+kt} \int_{-\infty}^{\infty} e^{-p^2} \, dp = e^{kt+x}.$$

Notice that $u(x,t)$ grows with time, which may seem to be in contradiction with the maximum principle. However, thinking in terms of heat conduction, we see that the initial temperature $u(x,0) = e^x$ is itself infinitely large at the far right end of the rod $x = +\infty$. So the temperature does not grow out of nowhere, but rather gets transferred from right to left with the "speed" k. Thus the initial exponential distribution of the temperature "travels" from right to left with the speed k as t grows. Compare this to the example in Strauss, where the initial temperature $u(x,0) = e^{-x}$ "travels" from left to right, since the initial temperature peaks at the far left end $x = -\infty$. □

In the above example we were able to compute the solution explicitly, however, the integral in (10.4) may be impossible to evaluate completely in terms of elementary functions for general initial data $\phi(x)$. Due to this, the answers for particular problems are usually written in terms of the error function in statistics,

$$\mathcal{E}\mathrm{rf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-p^2} \, dp.$$

Notice that $\mathcal{E}\mathrm{rf}(0) = 0$, and $\lim_{x\to\infty} \mathcal{E}\mathrm{rf}(x) = 1$. Using this function, we can rewrite the function $Q(x,t)$ given by (10.6), which solves the heat IVP with Heaviside initial data, as follows

$$Q(x,t) = \frac{1}{2} + \frac{1}{2}\mathcal{E}\mathrm{rf}\left(\frac{x}{\sqrt{4kt}}\right).$$

10.3 Conclusion

Using the notions of distribution and distributional derivative, we can make sense of the heat kernel $S(x,t)$ that has the Dirac δ function as its initial data. Comparing the expression of the heat kernel (10.3) with the density function of the normal (Gaussian) distribution, we saw that the solution formula (10.2) essentially weights the initial data by the bell-shaped curve $S(x,t)$, thus giving the contribution from the initial heat at different points towards the temperature at point x at time t.

11 Comparison of wave and heat equations

In the last several lectures we solved the initial value problems associated with the wave and heat equations on the whole line $x \in \mathbb{R}$. We would like to summarize the properties of the obtained solutions, and compare the propagation of waves to conduction of heat.

Recall that the solution to the wave IVP on the whole line

$$\begin{cases} u_{tt} - c^2 u_{xx} = 0, \\ u(x, 0) = \phi(x), \\ u_t(x, 0) = \psi(x), \end{cases} \tag{11.1}$$

is given by d'Alambert's formula

$$u(x, t) = \frac{1}{2}[\phi(x + ct) + \phi(x - ct)] + \frac{1}{2c}\int_{x-ct}^{x+ct} \psi(s)\, ds. \tag{11.2}$$

Most of the properties of this solution can be deduced from the solution formula, which can be understood fairly well, if one thinks in terms of the characteristic coordinates. This is how we arrived at the properties of finite speed of propagation, propagation of discontinuities of the data along the characteristics, and others.

On the other hand, the solution to the heat IVP on the whole line

$$\begin{cases} u_t - k u_{xx} = 0, \\ u(x, 0) = \phi(x), \end{cases} \tag{11.3}$$

is given by the formula

$$u(x, t) = \frac{1}{\sqrt{4\pi kt}} \int_{-\infty}^{\infty} e^{-(x-y)^2/4kt} \phi(y)\, dy. \tag{11.4}$$

We saw some of the properties of the solutions to the heat IVP, for example the smoothing property, in the case of the fundamental solution or the heat kernel

$$S(x, t) = \frac{1}{\sqrt{4\pi kt}} e^{-x^2/4kt}, \tag{11.5}$$

which had the Dirac delta function as its initial data. The solution u given by (11.4) can be written in terms of the heat kernel, and we use this to prove the properties for solutions to the general IVP (11.3). In terms of the heat kernel the solution is given by

$$u(x, t) = \int_{-\infty}^{\infty} S(x - y, t)\phi(y)\, dy = \int_{-\infty}^{\infty} S(z, t)\phi(x - z)\, dz,$$

where we made the change of variables $z = x - y$ to arrive at the last integral. Making a further change of variables $p = z/\sqrt{kt}$, the above can be written as

$$u(x, t) = \frac{1}{\sqrt{4\pi}} \int_{-\infty}^{\infty} e^{-p^2/4} \phi(x - p\sqrt{kt})\, dp. \tag{11.6}$$

This last form of the solution will be handy when proving the smoothing property of the heat equation, the precise statement of which is contained in the following.

Theorem 11.1. *Let $\phi(x)$ be a bounded continuous function for $-\infty < x < \infty$. Then (11.4) defines an infinitely differentiable function $u(x, t)$ for all $x \in \mathbb{R}$ and $t > 0$, which satisfies the heat equation, and* $\lim_{t \to 0+} u(x, t) = \phi(x)$, $\forall x \in \mathbb{R}$.

The proof is rather straightforward, and amounts to pushing the derivatives of $u(x,t)$ onto the heat kernel inside the integral. All one needs to guarantee for this procedure to go through is the uniform convergence of the resulting improper integrals. Let us first take a look at the solution itself given by (11.4). Notice that using the form in (11.6), we have

$$|u(x,t)| \leq \frac{1}{\sqrt{4\pi}} \int_{-\infty}^{\infty} \left| e^{-p^2/4} \phi(x - p\sqrt{kt}) \right| dp \leq \frac{1}{\sqrt{4\pi}} (\max|\phi|) \int_{-\infty}^{\infty} e^{-p^2/4}\, dp = \max|\phi|,$$

which shows that u, given by the improper integral, is well-defined, since ϕ is bounded. One can also see the maximum principle in the above inequality. We will use similar logic to show that the improper integrals appearing in the derivatives of u converge uniformly in x and t.

Notice that formally

$$\frac{\partial u}{\partial x} = \int_{-\infty}^{\infty} \frac{\partial S}{\partial x}(x-y,t)\phi(y)\, dy. \tag{11.7}$$

To make this rigorous, one must prove the uniform convergence of the integral. For this, we use expression (11.5) for the heat kernel to write

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{\partial S}{\partial x}(x-y,t)\phi(y)\, dy &= \frac{1}{\sqrt{4\pi kt}} \int_{-\infty}^{\infty} \frac{\partial}{\partial x} \left[e^{-(x-y)^2/4kt} \right] \phi(y)\, dy \\ &= -\frac{1}{\sqrt{4\pi kt}} \int_{-\infty}^{\infty} \frac{x-y}{2kt} e^{-(x-y)^2/4kt} \phi(y)\, dy. \end{aligned}$$

Making the change of variables $p = (x-y)/\sqrt{kt}$ in the above integral, we get

$$\int_{-\infty}^{\infty} \frac{\partial S}{\partial x}(x-y,t)\phi(y)\, dy = \frac{1}{4\sqrt{\pi kt}} \int_{-\infty}^{\infty} p e^{-p^2/4} \phi(x - p\sqrt{kt})\, dp \leq \frac{c}{\sqrt{t}} (\max|\phi|) \int_{-\infty}^{\infty} |p| e^{-p^2/4}\, dp,$$

where $c = 1/(4\sqrt{\pi k})$ is a constant. The last integral is finite, so the integral in the formal derivative (11.7) converges uniformly and absolutely for all $x \in \mathbb{R}$ and $t > \epsilon > 0$, where ϵ can be taken arbitrarily small. So the derivative $u_x = \partial u / \partial x$ exists and is given by (11.7).

The above argument works for the t derivative, and all the higher order derivatives as well, since for the n^{th} order derivatives one will end up with the integral $\int_{-\infty}^{\infty} |p|^n e^{-p^2/4}\, dp$, which is finite for all $n \in \mathbb{N}$. This proves the infinite differentiability of the solution, even though the initial data is only continuous.

We have already seen that u given by (11.4) solves the heat equation, due to the invariance properties. It then only remains to prove that $\lim_{t \to 0+} u(x,t) = \phi(x)$, $\forall x$. Recall that our previous proofs of this used the derivative of $\phi(x)$, or the language of distributions to employ the Dirac δ, where we assumed that ϕ is a test function, i.e. infinitely differentiable with compact support. To prove that $u(x,t)$ satisfies the initial condition in (11.3) in the case of continuous initial data ϕ as well, one can either use a *density argument*, in which $\phi(x)$ is uniformly approximated by smooth functions, and make a use of our earlier proofs, or provide a direct proof. The basic idea behind the direct proof is given next. We need to show that the difference $u(x,t) - \phi(x)$ becomes arbitrarily small when $t \to 0+$. First notice that

$$u(x,t) - \phi(x) = \int_{-\infty}^{\infty} S(x-y,t)[\phi(y) - \phi(x)]\, dy = \frac{1}{\sqrt{4\pi}} \int_{-\infty}^{\infty} e^{-p^2/4} [\phi(x - p\sqrt{kt}) - \phi(x)]\, dp, \tag{11.8}$$

where we used the same change of variables as before, $p = (x-y)/\sqrt{kt}$. To see that the last integral becomes arbitrarily small as t goes to zero, notice that if $p\sqrt{kt}$ is small, then $|\phi(x - p\sqrt{kt}) - \phi(x)|$ is small due to the continuity of ϕ, and the rest of the integral is finite. Otherwise, when $p\sqrt{kt}$ is large, then p is large, and the exponential in the integral becomes arbitrarily small, while the ϕ term is bounded. Thus, one estimates the above integral by breaking it into the following two integrals

$$\begin{aligned} &\frac{1}{\sqrt{4\pi}} \int_{-\infty}^{\infty} e^{-p^2/4} [\phi(x - p\sqrt{kt}) - \phi(x)]\, dp \\ &\quad = \frac{1}{\sqrt{4\pi}} \int_{|p| < \delta/\sqrt{kt}} e^{-p^2/4} [\phi(x - p\sqrt{kt}) - \phi(x)]\, dp + \frac{1}{\sqrt{4\pi}} \int_{|p| \geq \delta/\sqrt{kt}} e^{-p^2/4} [\phi(x - p\sqrt{kt}) - \phi(x)]\, dp. \end{aligned}$$

For some small δ, the first integral is small due to the continuity of ϕ, while for arbitrarily small t the second integral is the tail of a converging improper integral, and is hence small. You should try to fill in the rigorous details. This completes the proof of Theorem 11.1.

It turns out, that the result in the above theorem can be proved even if the assumption of continuity of ϕ is relaxed to piecewise continuity. One then has the following.

Theorem 11.2. *Let $\phi(x)$ be a bounded piecewise-continuous function for $-\infty < x < \infty$. Then* (11.4) *defines an infinitely differentiable function $u(x,t)$ for all $x \in \mathbb{R}$ and $t > 0$, which satisfies the heat equation, and*

$$\lim_{t\to 0+} u(x,t) = \frac{1}{2}[\phi(x+) + \phi(x-)], \qquad \text{for all } x \in \mathbb{R}, \tag{11.9}$$

where $\phi(x+)$ and $\phi(x-)$ stand for the right hand side and left hand side limits of ϕ at x.

Of course the fact that ϕ has jump discontinuities will not effect the convergence of the improper integrals encountered in the proof of Theorem 11.1. To see why (11.9) holds, notice that the integral in the right hand side of (11.6) can be broken into integrals over positive and negative half-lines

$$u(x,t) = \frac{1}{\sqrt{4\pi}} \int_{-\infty}^{0} e^{-p^2/4}\phi(x - p\sqrt{kt})\, dp + \frac{1}{\sqrt{4\pi}} \int_{0}^{\infty} e^{-p^2/4}\phi(x - p\sqrt{kt})\, dp.$$

Then, since $p < 0$ in the first integral, $\phi(x - p\sqrt{kt})$ goes to $\phi(x+)$ as $t \to 0+$, while it goes to $\phi(x-)$ in the second integral, due to p being positive. So one can make the obvious changes in the proof of the previous theorem to show (11.9). This curious fact is one of the reasons why some people prefer to define the value of the Heaviside step function at $x = 0$ to be $H(0) = \frac{1}{2}$. Then one has

$$\lim_{t\to 0+} Q(x,t) = H(x) \qquad \text{for all } x \in \mathbb{R} \quad (\text{including } x = 0!),$$

where $Q(x,t)$ was the solution arising from the initial data given by $H(x)$.

11.1 Comparison of wave to heat

We now summarize and compare the fundamental properties of the wave and heat equations in the table below. Brief discussion of each of the properties will follow.

	Property	**Wave** ($u_{tt} - c^2 u_{xx} = 0$)	**Heat** ($u_t - k u_{xx} = 0$)
(i)	Speed of propagation	Finite (speed $\leq c$)	Infinite
(ii)	Singularities for $t > 0$	Transported along characteristics (speed $= c$)	Lost immediately
(iii)	Well-posed for $t > 0$	Yes	Yes (for bounded solutions)
(iv)	Well-posed for $t < 0$	Yes	No
(v)	Maximum principle	No	Yes
(vi)	Behaviour as $t \to \infty$	Does not decay	Decays to zero (if ϕ is integrable)
(vii)	Information	Transported	Lost gradually

Let us now recall why each of the properties listed in the table holds or does not for each equation.

(i) Finite speed of propagation for the wave equation is immediately seen from d'Alambert's formula (11.2).

The infinite speed of propagation for the heat equation was seen in the example of the heat kernel, which is strictly positive for all $x \in \mathbb{R}$ for $t > 0$, but has Dirac δ function as its initial data, and hence is zero for all $x \neq 0$ initially.

(ii) We saw in the box-wave (initial displacement in the form of a box, no initial velocity) and the "hammer blow" (no initial displacement, initial box-shaped velocity) that singularities are preserved and are transported along the characteristics. The same is seen from (11.2).

For the heat equation we saw in the last section that the solution (11.4) is infinitely differentiable even for piecewise continuous initial data (this is true for even weaker conditions on ϕ).

(iii) Well-posedness for the wave IVP is seen immediately from d'Alambert's formula.

In the case of the heat equation, we proved uniqueness and stability using either the maximum principle, or alternatively, the energy method. Existence follows from our construction of the explicit solution (11.4).

(iv) For the wave equation, this follows from the invariance under time reversion. Indeed, if $u(x,t)$ is a solution, then so is $u(x,-t)$, which has data $(\phi(x), -\psi(x))$.

If we reverse the time in the heat equation, we get $u_t + ku_{xx} = 0, t > 0$. One can solve this equation in much the same way as the heat equation, and due to the symmetry in t, will get the solution

$$u(x,t) = \frac{1}{\sqrt{4\pi kt}} \int_{-\infty}^{\infty} e^{(x-y)^2/4kt} \phi(y)\, dy,$$

which diverges for all $x \in \mathbb{R}$ (unless ϕ decays to zero faster than e^{-p^2}). So the heat equation is not well-posed *backward in time.* This makes physical sense as well, since the processes described by the heat equation, namely diffusion, heat flow and random motion are irreversible processes.

(v) The fact that there is no maximum principle for the wave equation is apparent from the "hammer blow" example, where the solution was everywhere zero initially, but due to the nonzero initial velocity, had nonzero displacement for any time $t > 0$. For the heat equation, the maximum principle was proved rigorously in a previous lecture.

(vi) We saw that the energy is conserved for the wave equation, so the solutions do not decay. We also saw this in the box-wave example, in which the initial box-shaped data split into two box-shaped waves of half the height that traveled in opposite directions without changing the shape.

For the heat equation, the decay is seen from formula (11.4), since $S(x-y,t) \to 0$ as $t \to \infty$, and the integral will be bounded if ϕ is integrable. Notice that in the example we considered in the last lecture, with $\phi(x) = e^x$, the solution did not decay, but rather "traveled" from right to left. This was due to ϕ being non-integrable.

(vii) The fact that information is transported by the solutions of the wave equation is seen from the fact that the initial data is propagated along the characteristics. So the information will travel along the characteristics as well.

In the case of the heat equation, the information is gradually lost, which can be seen from the graph of a typical solution (think of the heat kernel). The heat from the higher temperatures gets dissipated and after a while it is not clear what the original temperatures were.

11.2 Conclusion

Although the wave and heat equations are both second order linear constant coefficient PDEs, their respective solutions posses very different properties. By now we have learned how to solve the initial value problems on the whole line for both of these equations, and understood these solutions in terms of the physics behind the corresponding problems. We also saw that the properties of the solutions of the respective equations correspond to our intuition for each of the physical phenomena described by the equations.